AFRL-VS-HA-TR-2004-1039

DOCUMENTATION OF SCALABLE INCREMENTAL MM5V3 4D-VAR SYSTEM

Xiaolei Zou
D. X. Zhang

Department of Meteorology
Florida State University
404 Love Building
Tallahassee, FL 32306-4520

14 April 2004

Final Report

20040730 034

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

AIR FORCE RESEARCH LABORATORY
Space Vehicles Directorate
29 Randolph Rd
AIR FORCE MATERIEL COMMAND
Hanscom AFB, MA 01731-3010

This technical report has been reviewed and is approved for publication.

/Signed/
FRANK RUGGIERO
Contract Manager

/Signed/
ROBERT BELAND
Branch Chief

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
14-04-2004

2. REPORT TYPE
Scientific, Final

3. DATES COVERED *(From - To)*
28 Sept 2001 - 31 Dec 2003

4. TITLE AND SUBTITLE
Documentation of Scalable Incremental MM5v3 4D-Var System

5a. CONTRACT NUMBER
F19628-01-C-0069

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER
NAVY

5e. TASK NUMBER
0T

5f. WORK UNIT NUMBER
01

6. AUTHOR(S)
Xiaolei Zou and D. X. Zhang

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Department of Meteorology
Florida State University
404 Love Building
Tallahassee, FL. 32306-4166

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory
AFRL/VSBYA
29 Randolph Road
Hanscom, AFB, MA 01731-3010

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-VS-HA-TR-2004-1039

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for Public Release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This report documents the changes made to the Penn State University/National Center for Atmospheric Research third revision of the Fifth Generation Mesoscale Model (MM5v3) four dimensional variational (4D-Var) analysis system to allow the execution of an incremental driver for the system. The incremental driver allows the user to select resolution or choice of physics that are used in the minimization of the cost function that are different than those used to describe the baisc state. In practice this would mean either a coarser resolution model grid or less computationally intensive physics routines would be chosen for the cost function minimization while a finer resolution model grid or more computationally intensive phyiscs for the basic state. The effect would be to significantly reduce the computations needed to complete the analysis. This report presents examples of the use of the incremental driver for the Hurricane Isaac case and compares the results to the original driver that uses consistent resolution and physics for the minimization and basic state computations. The results show that the use of less computationally complex physics in the minimization does not reduce the accuracy of the analysis or time to convergence. Using coarser resolution in the minimization does not reduce the accuracy of the analysis but does appear to increase the time to convergence.

15. SUBJECT TERMS
Mesoscale Models, Four-Dimensional Variational Analysis, Numerical Weather Prediction, Model Initialization, Incremental Four-Dimensional Variational Analysis

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Xiaolei Zou
Uunclassified	Unclassified	Unclassified	SAR	20	19b. TELEPHONE NUMBER *(Include area code)* (850) 644-6025

Contents

Figures

Tables

1. OVERVIEW

An incremental Four-Dimensional Variational (4D-Var) analysis system has been developed based on the scalable, third version of Penn State University/National Center for Atmospheric Research Fifth Generation Mesoscale Model (MM5v3) 4D-Var analysis software (Ruggiero et al. 2002). The above software consists of the normal 4D-Var non-incremental driver, an incremental 4D-Var driver, Tangent-linear Model (TLM) driver and adjoint (ADJ) driver. All are optimized to run on Massively Parallel Processors (MPP), distributed memory computers.

The incremental 4D-Var was developed to reduce the computational cost by providing flexibility on model resolution and physics. It follows the procedure outlined by Courtier et al. (1994). The incremental 4D-Var uses the non-linear model with higher resolution and/or more complicated surface and physics options to define an accurate basic state (trajectory). The tangent linear and adjoint model with a coarser resolution and/or simpler physical options are used to assimilate observations. For more general information on 4D-Var please see Zou et al. (1997; 1998).

Generally, the 4D-Var cost function can be defined as:

$$J(x(t_0)) = \tfrac{1}{2}\{x(t_0) - x_b\}^T B^{-1}\{x(t_0) - x_b\} + \tfrac{1}{2}\sum_{i=0}^{N}\{H_i x(t_i) - y_i\}^T O_i^{-1}\{H_i x(t_i) - y_i\}, \qquad (1)$$

where $\boldsymbol{x}$ is a vector of model variables. x_b is the background field in time t_0. B is the background error covariance matrix, $x(t_i)$ is the model forecast:

$$x(t_i) = M(t_i, t_0)x(t_0), \qquad (2)$$

where $M(t_i,t_0)$ is the non-linear model operator representing the integration from t_0 to t_i, H_i is the observation operator which maps model variables $\boldsymbol{x}$ to observation variable y:

$$y_i = H_i x(t_i), \qquad (3)$$

and O_i is the observational error covariance matrix.

The cost function that is minimized in an incremental 4D-Var is defined as

$$J(\delta x^n(t_0)) = \tfrac{1}{2}\{\delta x^n(t_0) + x^{n-1} - x_b\}^T B^{-1}\{\delta x^n(t_0) + x^{n-1} - x_b\}$$
$$+ \tfrac{1}{2}\sum_{i=0}^{N}\{y_i^{n-1} + H_i'\delta x^n(t_i) - y_i\}^T O_i^{-1}\{y_i^{n-1} + H_i'\delta x^n(t_i) - y_i\} \qquad (4)$$

in which

$$\delta x^n(t_i) = R(t_i, t_0)\delta x^n(t_0) \qquad (5)$$

and

$$y_i^{n-1} = H_i\{M(t_i,t_0)(x^{n-2} + \delta^{*n-1}x(t_0))\} = H_i\{M(t_i,t_0)(x^{n-1})\}, \tag{6}$$

where δx is the model perturbation from the basic state. R is the tangent linear model operator, the superscript n represents the nth basic state, and $\delta^{*n-1}x(t_0)$ is the minimization of the cost function in the $(n-1)$th step where $\delta^{*-1}x(t_0)=0$.

$$x^n = x^{n-1} + \delta^{*n}x(t_0) \qquad (x^0 = x_b, x^{-1} = x_b) \tag{7}$$

To implement the above incremental 4D-Var, two loops (outer and inner) are needed. In the outer loop, a non-linear model integration using (6) is carried out to generate the model trajectory. The inner loop uses the tangent linear (5) and adjoint model to minimize the cost function (4) to obtain an optimal analysis increment. Once the inner loop is completed, the optimal analysis increment is added to the model initial condition, which is used for generating the previous basic state, and a new basic state is created in the next outer loop by (7).

2. SYSTEM STRUCTURE

2.1 Non-linear Model

In the non-linear forward model all global symbols such as: subroutine names, function names, block data names, and common block names are self-enclosed. Calls from outside are limited to the main subroutine called "MM5_0." The non-linear model exchanges data with other subroutines by calling a subroutine that is described in section 2.4 as the interface.

The new global names in the nonlinear model are based on the old ones with suffix "_0" attached to each of them. For example, the subroutine "MM5" is changed to a subroutine named "MM5_0".

2.2 Tangent Linear and Adjoint Models

The tangent-linear model replaces the non-linear model in the 4D-Var minimization procedure. The basic states (trajectory) are read in at every time step in the tangent linear model integration. No change is applied to the adjoint model.

2.3 I/O System

Both of the non-linear models and the minimization part (including tangent-linear and adjoint models) have separate I/O systems, different I/O unit numbers and different files. Users should provide a set of files to each of them. The minimization part uses the unit numbers and file names in the previous system. The non-linear model uses new unit numbers, new file names, and new namelist names. For example, the MM5 initial input file name in the non-linear model is "MMINPUT_DOMAIN1_0." In the minimization part, the corresponding file name is "MMINPUT_DOMAIN1."

2.4 Interface of Data Exchange

Subroutine "EXCHGIO" is the only one to be used to exchange data (trajectory) between the non-linear model and the minimization part and to update the initial fields for the non-linear model. It is called by subroutine "MM5_0." An interpolation is applied when the non-linear model and the minimization part have different horizontal resolutions. When going from a high resolution to a lower resolution a nine-point weighting averaging is used and when going from low resolution to a higher resolution a bilinear interpolation is used. The same interpolation subroutine is used for inner-loop runs.

2.5 Structure of Source Code

The tangent linear model replaces the non-linear model in the minimization. The new non-linear model is organized independently. In the main program "FDVDRIVER," an outer loop is inserted into the old minimization loop in which the non-linear forward model is integrated to provide the trajectory for the next minimization (inner) loop. The analysis increments (results of the minimization with inner loops) are then used to update the non-linear model initial fields. Users can adjust the numbers of outer and inner loops for a desired balance between accuracy and computational cost.

3. SYSTEM BUILDING

3.1. Configuration File and Makefile

There are two separate model integration systems (non-linear model and tangent linear/adjoint model) in the entire incremental 4D-Var system. Each of them can have their own resolution and physics options, and each have their own configuration file to set parameters for them. File "configure.user" contains the user-specified minimization parameter settings and file "configure_0.user" contains the settings for the non-linear model. The Makefile has two threads to make the two model systems and then build a 4D-Var executable with all objective libraries created.

3.2 Namelist Files

By editing the namelist files, users can change the domain resolution, planetary boundary layer physics, cumulus parameterization, and other settings. The non-linear model and the minimization part have different namelist file names and different namelist names. For example, the namelist file "./Run/oparam_0" for the non-linear model may contain "&OPARAM_0 TIMAX = 45., TISTEP = 10., ..., &END"; while the namelist file "./Run/oparam" for the minimization part may contain "&OPARAM TIMAX = 45., TISTEP = 30., ..., &END".

3.3 Input Data Preparation

The whole system needs two sets of files as initial input. For the same resolution-setting, the two sets are the same. For different resolution settings, the input data have the same domain coverage, but different grid distances. Currently, the 4D-Var system does not provide the mechanics to interpolate the initial input data internally. Users can easily generate two sets of these input data by using MM5v3 preprocessing software (see Dudhia et al. 2004).

A direct observation file is necessary to calculate the scaling and weighting when assimilating more than one type of observations. When making the 4D-Var executable, an executable for the direct observations is created simultaneously. Submitting the script file "dirobs.deck" will generate the direct observation file "DIROBS_DOMAIN1" in sub-directory "Run".

3.4 Job Deck

The job deck is similar to that for non-incremental 4D-Var. All namelist files are created by the shell scripts. Careful attention should be placed on the values of TIMAX and TISTEP.

4. SYSTEM APPLICATION

4.1 Different Physics Options

The options to select microphysics (IMPHYS), cumulus parameterization (ICUPA), planetary boundary layer (IBLTYP) etc. are controlled in the "configure.user" and "configure_0.user" files. After modifications are made, the user should clean out the old objective and library files and rebuild the executable. Also, the user can tune the parameters in namelist files "./Run/lparam" and "./Run/lparam_0" by editing the job deck.

4.2 Different Resolutions

To set different resolutions in the non-linear model and minimization part, the user should edit the configure.user files and job decks to have them ready. Different

physics options can be selected simultaneously in coarse domain and fine resolution domain. Executable rebuilding is necessary after the files "configure.user" or "configure_0.user" are modified.

4.3 Consistence of Settings

No validation checking is available for user specified parameters in the current system. Serious problems can occur with improper parameter settings. For example, the maximum forecast time of the non-linear model should be equal to or greater than the assimilation time used by the minimization part. It is the user's responsibility to make sure that all parameters are set consistently.

5. TEST EXPERIMENTS

The experiments presented here are for the case of Hurricane Issac in 2000 (Franklin et al. 2001). In the tests, the Bogus Data Assimilation (BDA) technique is used following the procedure outlined by Zou and Xiao (2000). Specifics on running the BDA technique are given in Appendix A. The inputs for constructing the bogus data for the hurricane vortex comes from the observations of the National Hurricane Center and are listed in Table 1. The time that the optimal analyses from the 4D-Var were generated is 26 September 2000 at 0000 UTC.

Table 1 Observations of Hurricane Isaac

NAME	DATE	TIME	PRESS (hPa)	Pout (hPa)	LOCATION	Rout (nm)	R (km)	MaxWind (kt)
ISAAC	20000926	00	970	1012	(17.9,-42.0)	180	88	90
		12	980	1013	(18.6,-43.9)	150	88	75
	20000927	00	977	1013	(19.6,-46.0)	150	77	80
		12	970	1013	(21.0,-48.1)	150	62	90
	20000928	00	970	1013	(22.8,-50.6)	150	96	100
		12	950	1012	(25.0,-52.9)	150	100	110
	20000929	00	948	1012	(28.0,-55.1)	150	113	115

5.1 Control Run (non-incremental 4DVAR)

The horizontal resolution is 30 km, with the domain size 49x49 and 27 vertical levels (half sigma level). The following is the physics options section in file "configure.user":

```
# IMPHYS - for explicit moisture schemes (array,integer)
```

```
IMPHYS = "2,1,1,1,1,1,1,1,1,1"
#                  - Dry,stable,warm rain,simple ice,mix phase,
#                  - 1  ,2     ,3        ,4         ,5
#                  - graupel(gsfc),graupel(reisner2),schultz
#                  -,6            ,7                ,8
MPHYSTBL = 0
#                  - 0=do not use look-up tables for moist
#                     physics
#                  - 1=use look-up tables for moist physics
#                   (currently only simple ice and mix phase
#                    are available)
#
# ICUPA - for cumulus schemes (array,integer)
#                  - None,Kuo,Grell,AS,FC,KF,BM - 1,2,3,4,5,6,7
ICUPA  = "2,1,1,1,1,1,1,1,1,1"
#
# IBLTYP - for planetary boundary layer (array,integer)
#                  - 0=no PBL fluxes,1=bulk,2=Blackadar,
#                    3=Burk-Thompson,4=Eta M-Y,5=MRF,
#                    6=Gayno-Seaman
IBLTYP  = "5,0,0,0,0,0,0,0,0,0"
#
# FRAD - for atmospheric radiation (integer)
#                  - Radiation cooling of atmosphere
#                    0=none,1=simple,2=cloud,3=ccm2,4=rrtm
FRAD = "1,0,0,0,0"
#
# ISOIL - for multi-layer soil temperature model (integer)
#                  - 0=no,1=yes (only works with IBLTYP=2,4,5,6)
#                    2=OSU land-surface scheme (IBLTYP=5)
ISOIL = 1
#
# ISHALLO (array,integer) - Shallow Convection Option
#                1=shallow convection,0=No shallow convection
ISHALLO  = "0,0,0,0,0,0,0,0,0,0"
```

The assimilation window is set to 30 minutes. A total of 30 iterations are been carried out in this experiment.

5.2 Incremental 4D-Var with the Same Settings in Nonlinear and Adjoint Models

Numerical results from the incremental 4D-Var experiment with the above settings are compared with those from the non-incremental 4D-Var (EXP1). The assimilation window is 30 minutes. The outer loop for updating the non-linear initial conditions is 30 and the inner loop for optimizing the cost function is 1. The total iteration number is 30.

5.3 Different Physics Options in Non-linear Model and Adjoint Model

This is EXP2. The configuration files configure.user and configure_0.user contain different settings. Here are the differences between the two files with ">" referring to file configure_0.user and "<" referring to configure.user.

```
18c18
< # 5. Options for making "./include/parame.incl"
---
> # 5. Options for making "./include_0/parame_0.incl"
42c42
< LIBINCLUDE = $(DEVTOP)/include
---
> LIBINCLUDE = $(DEVTOP)/include_0
265c265
< # 5. Options for making ./include/parame.incl
---
> # 5. Options for making ./include_0/parame_0.incl
306c306
< ICUPA  = "2,1,1,1,1,1,1,1,1,1"
---
> ICUPA  = "3,1,1,1,1,1,1,1,1,1"
312c312
< IBLTYP  = "5,0,0,0,0,0,0,0,0,0"
---
> IBLTYP  = "2,0,0,0,0,0,0,0,0,0"
```

The assimilation window is also 30 minutes. The outer loop for updating the non-linear initial conditions is 30 and the inner loop for optimizing the cost function is 1. Total number of iterations is 30.

5.4 Different Resolution – Same Physics Options in Non-linear and Adjoint Models

In this experiment (EXP3), the horizontal resolution for the non-linear model is reduced to 15 km and the domain size is 145x145. Here are the differences between the two files.

```
< # 5. Options for making "./include/parame.incl"
---
> # 5. Options for making "./include_0/parame_0.incl"
42c42
```

```
< LIBINCLUDE = $(DEVTOP)/include
---
> LIBINCLUDE = $(DEVTOP)/include_0
266c266
< # 5. Options for making ./include/parame.incl
---
> # 5. Options for making ./include_0/parame_0.incl
279,280c279,280
< MIX = 49
< MJX = 49
---
> MIX =145
> MJX =145
```

The assimilation window is 30 minutes. The outer loop for updating the nonlinear initial conditions is 4 and the inner loop for optimizing the cost function is 8. The total number of iterations is 32.

6. RESULTS

Output of the cost function for each of the four runs as a function of iteration is given in Figure 1. All four runs show reasonable convergence by the 20^{th} iteration. The EXP1 and EXP2 practically mimic each other exactly and converge faster than the Control run. EXP3 is the slowest to converge and also converges to a slightly larger cost function than the other three runs. Figure 2 contains the norm of the gradient for the four runs and shows similar behavior to that seen in Figure 1. The optimal initial sea-level pressure fields are shown in Figure 3. Although EXP3 has 32 iterations, the fields in Figure 3 show the result after 30 iterations. All the experiments have the similar initial sea level pressure fields. It seems that the different physics options and resolution have very little effect on the optimal results compared with the results using the same physics and resolution options.

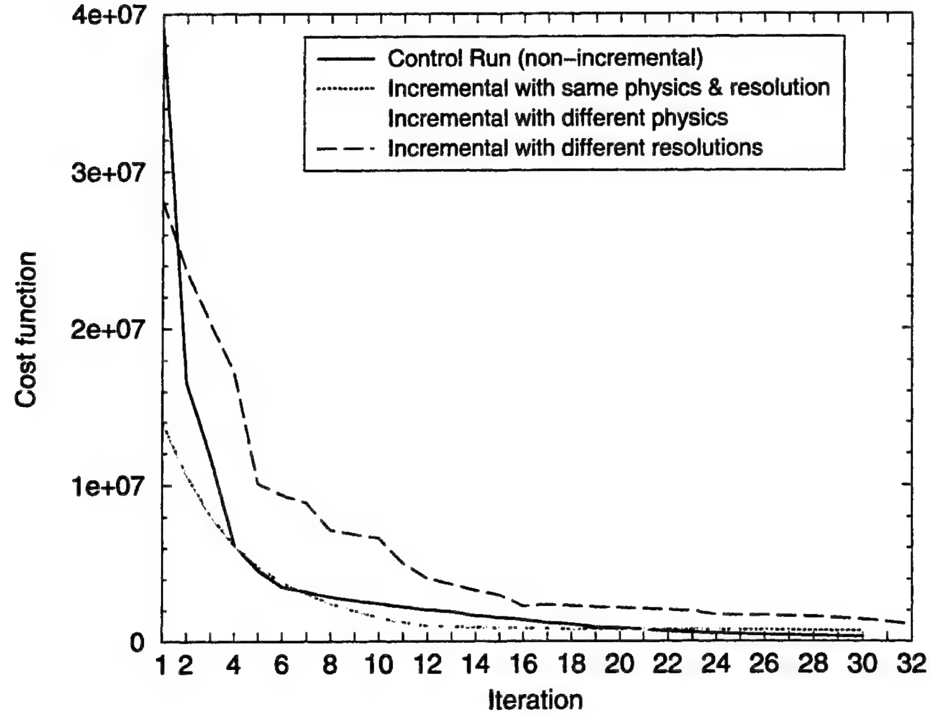

Figure 1 Total Cost Function at Each Iteration.

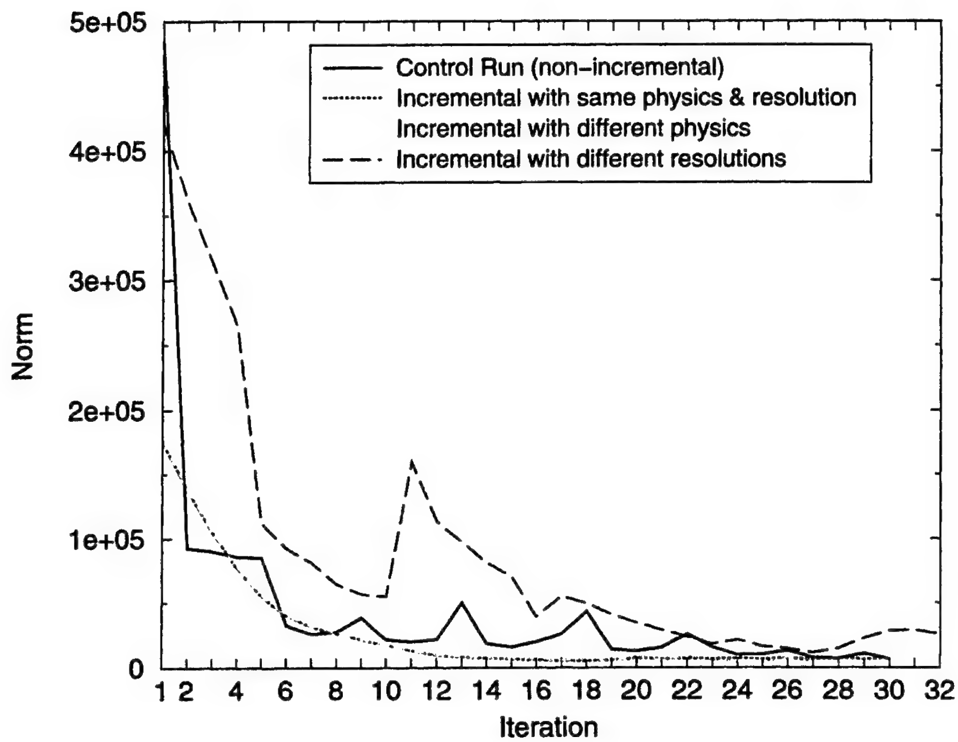

Figure 2 Evolution of the Norm of the Gradient as a Function of Iterations.

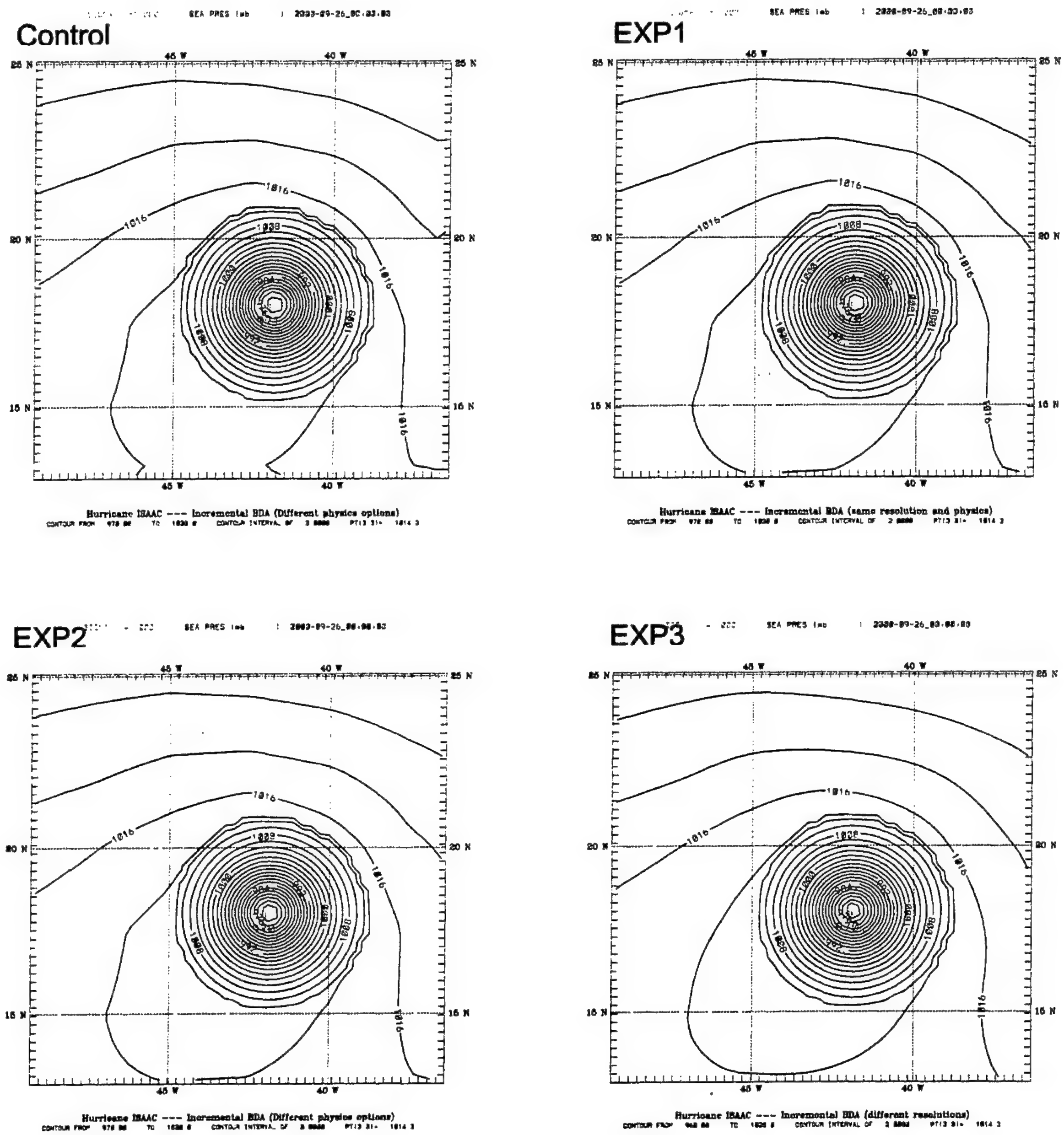

Figure 3 Optimized Sea-level Pressure (units in hPa).

REFERENCES

Courtier, P., J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteorol. Soc.*, **120**, 1367-1378.

Dudhia, J. D. Gill, K. Manning, W. Wang, and C. Bruyere, 2004: PSU/NCAR Mesoscale Modeling System Tutorial Class Notes and Users' Guide *(MM5 Modeling System Version 3)* http://www.mmm.ucar.edu/mm5/documents/tutorial-v3-notes-pdf/tutorial.cover.pdf.

Franklin, J. L., L. A. Avila, J. L. Beven, M. B. Lawrence, R. J. Pasch, S. R. Stewart, 2001, Atlantic Hurricane Season of 2000, *Mon. Wea. Rev.,* **129**, 3037-3056

Ruggiero, F. H., D. Norquist, T. Nehrkorn, G. D. Modica, M. Cerniglia, J. G. Michalakes, and X. Zou, 2002: MM5v3-based 4DVAR: Current status and future plans. Preprints, 12th PSU/NCAR Mesoscale Model Users' Workshop, 24-25 June 2002, Boulder CO. http://www.mmm.ucar.edu/mm5/workshop/ws02/Ruggiero.pdf

Zou, X., and Q. Xiao, 2000: Studies on the initialization and simulation of a mature hurricane using a variational bogus data assimilation scheme. *J. Atmos. Sci.,* **57**, 836-860.

Zou, X., F. Vandengerghe, M. Pondeca, and Y.-H. Kuo, 1997: Introduction to adjoint techniques and the MM5 adjoint modeling system. NCAR Tech. Note NCAR/TN-435-STR, NCAR, P.O. Box 3000 Boulder, CO 80307.

Zou X., W. Huang, and Q. Xiao, 1998: A user's guide to the MM5 adjoint modeling system. NCAR Tech. Note NCAR/TN-437+IA, NCAR, P.O. Box 3000 Boulder, CO 80307.

APPENDIX

Implementation of the Bogus Data Assimilation Technique

The Bogus Data Assimilation system used here is similar to that described in Zou and Xiao (2000) and is based on the MM5 version 3 4D-Var system. It consists of three parts: (1) Generating the hurricane bogus data of the sea level pressure (SLP) field, (2) running 4D-Var to fit the model fields to the bogus SLP, and (3) merging the bogussed vortex to the MM5 initial conditions.

Bogus.exe will create a file of bogus sea-level pressure observations based on the Fujita vortex model:

$$P(r) = P_{\inf} - \frac{(P_{\inf} - P_c)}{\sqrt{1 + \left(\frac{r}{R_0}\right)^2}} \qquad \text{(A1)}$$

where p(r) is the sea level pressure at radial distance r, P_c is the central minimum pressure, R_0 is the radius of maximum gradient of the SLP multiplied by $\sqrt{2}$ and can be estimated by the following model described by Park and Zou (2004):

$$R_0 = 0.37 R_{34kt} + 1.9 \qquad \text{(A2)}$$

where R_{34kt} is 34kt wind radii. P_{inf} is the sea level pressure at infinity distance and can be found by solving Fujita's fomula to have P_{out} at $r=R_{out}$ for a given P_c and R_0. P_{out} and R_{out} are the sea level pressure and radius of the outer most closed isobar.

To compile and execute bogus.exe :

- Move to the directory containing the source code for bogus.exe

- Type "make" in this directory to see what you should do next

- When "make" is successful, executable bogus.exe will appear in this directory

- Type: % bogus.exe -Tbogus.nl -IMMINPUT_DOMAIN1 -OBOGUS_OBS

where bogus.nl is the namelist file that contains the parameters for computing the bogus vortex as described above. MMINPUT_DOMAIN1 is the MM5 model input field that the optimal initial conditions are being constructed for and BOGUS_OBS is the output file of bogus sea level pressure observations. Typically, when running the BDA scheme in the 4D-Var, a smaller domain than the one specified by the MMINPUT_DOMAIN1 is used. The domain is chosen to cover the area of the bogus vortex and is done to save computational expense. To merge the results of the 4D-Var with the original MMINPUT_DOMAIN1 file the program overlap is used.

To compile and execute overlap.exe :

- Overlap.exe will be created when running "make" for bogus.exe

- When "make" is successful, executable overlap.exe will appear in the same directory

- Type % overlap.exe -IMMINPUT_DOMAIN1 -NMMINPUT_DOMAIN2 -LLAST.0030 -OMMINPUT_DOMAIN1_opt

where MMINPUT_DOMAIN1 is the original MM5 initial conditions, MMINPUT_DOMAIN2 is the domain that the 4D-Var executed on, LAST.00XX is the last iteration file from the 4D-VAR and MMINPUT_DOMAIN1_opt is the output of the optimal initial conditions for which to run the forecast model.

Reference.

Park, K., and X. Zou, 2004: Toward developing an objective 4D-Var BDA scheme for hurricane initialization based on TPC observed parameters. *Mon. Wea. Rev.*, accepted for publication.